

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 101671995A  
Source: IPHO  
Date Processed by STIC: 1-27-05

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 01/27/2005

PATENT APPLICATION: US/10/671,995A

TIME: 14:52:47

Input Set : A:\BHPat.txt

Output Set: N:\CRF4\01272005\J671995A.raw

3 <110> APPLICANT: BioHesion, Inc  
 4 Woodbury, Richard G  
 5 deVos, Theo  
 6 Irani, Meher  
 8 <120> TITLE OF INVENTION: Recombinant Fusion Proteins with High Affinity Binding to  
 Gold  
 9 and Applications Thereof  
 11 <130> FILE REFERENCE: Fusionpatent.doc  
 13 <140> CURRENT APPLICATION NUMBER: US 10/671,995A  
 14 <141> CURRENT FILING DATE: 2003-09-26  
 16 <160> NUMBER OF SEQ ID NOS: 8  
 18 <210> SEQ ID NO: 1  
 19 <211> LENGTH: 5454  
 20 <212> TYPE: DNA  
 21 <213> ORGANISM: Escherichia coli  
 24 <220> FEATURE:  
 25 <223> OTHER INFORMATION: Expression plasmid pPA-GBP, for His6-protein A-GBP. Vector is  
 26 pQE-80L (Qiagen). CDS for Protein A (nucleotides 160-528) is from  
 27 pEZZ18 (Amersham). CDS for GBP (nucleotides 565-858) is from  
 28 pSB3053 from (Brown, Nat. Biotechnol. 15:269-272, 1997).  
 31 <220> FEATURE:  
 32 <221> NAME/KEY: CDS  
 33 <222> LOCATION: (115)..(885)  
 34 <223> OTHER INFORMATION: CDS for fusion protein His6-protein A-GBP consisting of 257  
 amino  
 35 acid residues.  
 37 <400> SEQUENCE: 1  
 38 ctcgagaaat cataaaaaat ttatttgctt tgtgagcggg taacaattat aatagattca 60  
 40 attgtgagcg gataacaatt tcacacagaa ttcattaaag aggagaaatt aact atg 117  
 41 Met  
 42 1  
 44 aga gga tgc cat cac cat cac cat cac gga tcc ggt tct ggt gcg caa 165  
 45 Arg Gly Ser His His His His His His Gly Ser Gly Ser Gly Ala Gln  
 46 5 10 15  
 48 cac gat gaa gcc gta gac aac aaa ttc aac aaa gaa caa caa aac gcg 213  
 49 His Asp Glu Ala Val Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala  
 50 20 25 30  
 52 ttc tat gag atc tta cat tta cct aac tta aac gaa gaa caa cga aac 261  
 53 Phe Tyr Glu Ile Leu His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn  
 54 35 40 45  
 56 gcc ttc atc caa agt tta aaa gat gac cca agc caa agc gct aac ctt 309  
 57 Ala Phe Ile Gln Ser Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu  
 58 50 55 60 65  
 60 tta gca gaa gct aaa aag cta aat gat gct cag gcg ccg aaa gta gac 357

61 Leu Ala Glu Ala Lys Lys Leu Asn Asp Ala Gln Ala Pro Lys Val Asp

## RAW SEQUENCE LISTING

DATE: 01/27/2005

PATENT APPLICATION: US/10/671,995A

TIME: 14:52:47

Input Set : A:\BHPat.txt

Output Set: N:\CRF4\01272005\J671995A.raw

62		70		75		80	
64	aac aaa ttc	aac aaa gaa caa caa aac gcg ttc tat gag atc tta cat	405				
65	Asn Lys Phe	Asn Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile Leu His					
66		85	90	95			
68	tta cct aac	tta aac gaa gaa caa cga aac gcc ttc atc caa agt tta	453				
69	Leu Pro Asn	Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser Leu					
70		100	105	110			
72	aaa gat gac	cca agc caa agc gct aac ctt tta gca gaa gct aaa aag	501				
73	Lys Asp Asp	Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys Lys					
74		115	120	125			
76	cta aat gat	gct cag gcg ccg aaa gta gac gcg aat tcg agc tct ggt	549				
77	Leu Asn Asp	Ala Gln Ala Pro Lys Val Asp Ala Asn Ser Ser Ser Gly					
78	130	135	140	145			
80	agt ggc aat	ggt cat atg cat gga aaa act cag gca acc agc ggg act	597				
81	Ser Gly Asn	Gly His Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr					
82		150	155	160			
84	atc cag agc	atg cat gga aaa act cag gca acc agc ggg act atc cag	645				
85	Ile Gln Ser	Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln					
86		165	170	175			
88	agc atg cat	gga aaa act cag gca acc agc ggg act atc cag agc atg	693				
89	Ser Met His	Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met					
90		180	185	190			
92	cat gga aaa	act cag gca acc agc ggg act atc cag agc atg cat gga	741				
93	His Gly Lys	Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly					
94		195	200	205			
96	aaa act cag	gca acc agc ggg act atc cag agc atg cat gga aaa act	789				
97	Lys Thr Gln	Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr					
98	210	215	220	225			
100	cag gca acc	agc ggg act atc cag agc atg cat gga aaa att cag gca	837				
101	Gln Ala Thr	Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile Gln Ala					
102		230	235	240			
104	acc agc ggg	act atc cag agc atg cat gct ctg tcc ctc gag ggt ccg	885				
105	Thr Ser Gly	Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu Gly Pro					
106		245	250	255			
108	taataagctt	aattagctga gcttggactc ctgttgatag atccagtaat gacctcagaa	945				
110	ctccatctgg	atttggtcag aacgctcggg tgccgccggg cgttttttat tgggtgagaat	1005				
112	ccaagctagc	ttggcgagat ttccaggagc taaggaagct aaaatggaga aaaaaatcac	1065				
114	tggatatacc	accgttgata tatcccaatg gcatcgtaaa gaacattttg aggcatttca	1125				
116	gtcagttgct	caatgtacct ataaccagac cgttcagctg gatattacgg cctttttaaa	1185				
118	gaccgtaaag	aaaaataagc acaagtttta tccggccttt attcacattc ttgcccgcct	1245				
120	gatgaatgct	catccggaat ttcgtatggc aatgaaagac ggtgagctgg tgatatggga	1305				
122	tagtgttcac	ccttgttaca ccgttttcca tgagcaaact gaaacgtttt catcgctctg	1365				
124	gagtgaatac	cacgacgatt tccggcagtt tctacacata tattcgcaag atgtggcgtg	1425				
126	ttacggtgaa	aacctggcct atttccctaa aggggtttatt gagaatatgt ttttcgtctc	1485				
128	agccaatccc	tgggtgagtt tcaccagttt tgatttaaac gtggccaata tggacaactt	1545				
130	cttcgcccc	gttttcacca tgggcaaata ttatacgcaa ggcgacaagg tgctgatgcc	1605				
132	gctggcgatt	caggttcatc atgccgtttg tgatggcttc catgtcggca gaatgcttaa	1665				
134	tgaattacaa	cagtactgcg atgagtggca gggcggggcg taattttttt aaggcagtta	1725				
136	ttggtgcct	taaacgcctg gggtaatgac tctctagctt gaggcatcaa ataaaacgaa	1785				

## RAW SEQUENCE LISTING

DATE: 01/27/2005

PATENT APPLICATION: US/10/671,995A

TIME: 14:52:47

Input Set : A:\BHPat.txt

Output Set: N:\CRF4\01272005\J671995A.raw

138	aggctcagtc	gaaagactgg	gcctttcggt	ttatctgttg	tttgtcggtg	aacgctctcc	1845
140	tgagtaggac	aaatccgccc	tctagattac	gtgcagtcga	tgataagctg	tcaaaccatga	1905
142	gaattgtgcc	taatgagtga	gctaacttac	attaattgcg	ttgcgctcac	tgcccgcctt	1965
144	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	ggccaacgcg	cggggagagg	2025
146	cggtttgctg	attgggcgcc	agggtgggtt	ttcttttcac	cagtgaagacg	ggcaacagct	2085
148	gattgccctt	caccgcctgg	ccctgagaga	gttgacgcaa	gcggtccacg	ctggtttgcc	2145
150	ccagcaggcg	aaaatcctgt	ttgatgggtg	ttaacggcgg	gatataacat	gagctgtctt	2205
152	cggatcgtc	gtatccact	accgagatat	ccgcaccaac	gcgcagcccg	gactcggtaa	2265
154	tggcgcgcac	tgccccagc	gccatctgat	cgttggcaac	cagcatcgca	gtgggaacga	2325
156	tgcctcatt	cagcatttgc	atggtttgtt	gaaaaccgga	catggcactc	cagtcgcctt	2385
158	cccgttccgc	tatcggtcga	atttgattgc	gagtgaagata	tttatgccag	ccagccagac	2445
160	gcagacgcgc	cgagacagaa	cttaattggc	ccgctaacag	cgcgatttgc	tggtgaccca	2505
162	atgcgaccag	atgctccacg	cccagtcgcg	taccgtcttc	atgggagaaa	ataatactgt	2565
164	tgatgggtgt	ctggtcagag	acatcaagaa	ataacgcccg	aacattagtg	caggcagctt	2625
166	ccacagcaat	ggcatcctgg	tcatccagcg	gatagttaat	gatcagccca	ctgacgcgtt	2685
168	gcgcgagaag	attgtgcacc	gccgctttac	aggcttcgac	gccgcttcgt	tctaccatcg	2745
170	acaccaccac	gctggcacc	agttgatcgg	cgcgagattt	aatcgccgcg	acaatttgcg	2805
172	acggcgcggtg	cagggccaga	ctggaggtgg	caacgccaat	cagcaacgac	tgtttgccc	2865
174	ccagttgttg	tgccacgcgg	ttgggaatgt	aattcagctc	cgccatcgcc	gcttccactt	2925
176	tttcccgctg	tttcgcagaa	acgtggctgg	cctggttcac	cacgcgggaa	acggtctgat	2985
178	aagagacacc	ggcatactct	gcgacatcgt	ataacgttac	tggtttcaca	ttcaccaccc	3045
180	tgaattgact	ctcttcgggg	cgctatcatg	ccataccgcg	aaaggttttg	caccattcga	3105
182	tggtgtcggg	atttcgggca	gcgttgggtc	ctggccacgg	gtgcgcacga	tctagagctg	3165
184	cctcgcgctg	ttcgggtgatg	acgggtgaaa	cctctgacac	atgcagctcc	cggagacggt	3225
186	cacagcttgt	ctgtaagcgg	atgccgggag	cagacaagcc	cgtcagggcg	cgtcagcggg	3285
188	tggttggcggg	tgctcggggcg	cagccatgac	ccagtcacgt	agcgatagcg	gagtgatatac	3345
190	tggtttaact	atgcggcatc	agagcagatt	gtactgagag	tgccaccat	gcggtgtgaa	3405
192	ataccgcaca	gatgcgtaag	gagaaaatac	cgcatcaggc	gctcttcgcg	ttcctcgcctc	3465
194	actgactcgc	tgcgctcggg	cggttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	3525
196	gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	3585
198	cagcaaaagg	ccaggaaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	3645
200	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgcagga	3705
202	ctataaagat	accaggcggt	ttccccctgga	agctccctcg	tgcgctctcc	tggtccgacc	3765
204	ctgcgcctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcggtgg	gctttctcat	3825
206	agctcacgct	gtaggtatct	cagttcgggtg	taggtcgttc	gctccaagct	gggctgtgtg	3885
208	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagctc	3945
210	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	4005
212	gcgaggtatg	taggcgggtg	tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	4065
214	agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	4125
216	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtgggttttt	tggttgcaag	4185
218	cagcagatta	cgcgacagaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacggg	4245
220	tctgacgctc	agtggaacga	aaactcacgt	taagggattt	tggtcatgag	attatcaaaa	4305
222	aggatcttca	cctagatcct	tttaaatata	aatgaagtt	ttaaatcaat	ctaaagtata	4365
224	tatgagtaaa	cttgggtctga	cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	4425
226	atctgtctat	ttcggttcac	catagttgcc	tgactccccg	tcgtgtagat	aactacgata	4485
228	cgggagggct	taccatctgg	ccccagtgct	gcaatgatac	cgcgagaccc	acgctcaccg	4545
230	gctccagatt	tatcagcaat	aaaccagcca	gccggaagg	ccgagcgag	aagtggctct	4605
232	gcaactttat	ccgcctccat	ccagctctatt	aattgttgcc	gggaagctag	agtaagtagt	4665
234	tcgccagttta	atagtttgcg	caacgttggt	gccattgcta	caggcatcgt	ggtgtcacgc	4725

## RAW SEQUENCE LISTING

DATE: 01/27/2005

PATENT APPLICATION: US/10/671,995A

TIME: 14:52:47

Input Set : A:\BHPat.txt

Output Set: N:\CRF4\01272005\J671995A.raw

```

236 tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 4785
238 tcccccatgt tgtgcaaaaa agcggttagc tccttcgggc ctccgatcgt tgtcagaagt 4845
240 aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 4905
242 atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 4965
244 tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtaaa tacgggataa taccgcgcca 5025
246 catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 5085
248 aggatcttac cgctgttgag atccagttcg atgtaaccga ctcgtgcacc caactgatct 5145
250 tcagcatctt ttactttcac cagcgtttct gggtagacaa aaacaggaag gcaaaatgcc 5205
252 gcaaaaaagg gaataaggcg gacacggaaa tgttgaatac tcatactctt cctttttcaa 5265
254 tattattgaa gcattttca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 5325
256 tagaaaaata aacaaatagg ggttcgcgc acatttcccc gaaaagtgcc acctgacgtc 5385
258 taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac gaggcccttt 5445
260 cgtcttcac 5454
264 <210> SEQ ID NO: 2
265 <211> LENGTH: 5448
266 <212> TYPE: DNA
267 <213> ORGANISM: Escherichia coli
270 <220> FEATURE:
271 <223> OTHER INFORMATION: Expression plasmid pStreptavidin-GBP for His6-streptavidin-
GBP.
272 pQE-80L vector(Qiagen). CDS core streptavidin(169-531), (Chilkoti
273 et al., Proc Natl Acad Sci USA 92:1754-1758, 1995). CDS GBP
274 (559-852), pSB3053, (Brown, Nat. Biotechnol. 15:269-272, 1997).
277 <220> FEATURE:
278 <221> NAME/KEY: CDS
279 <222> LOCATION: (115)..(879)
280 <223> OTHER INFORMATION: CDS for fusion protein His6-streptavidin-GBP consisting of
255 amino acid residues.
281 amino acid residues.
283 <400> SEQUENCE: 2
284 ctcgagaaat cataaaaaat ttatttgctt tgtgagcgga taacaattat aatagattca 60
286 attgtgagcg gataacaatt tcacacagaa ttcattaaag aggagaaatt aact atg 117
287 Met
288 1
290 aga gga tcg cat cac cat cac cat cac gga tcc ggt tct ggt ggc cat 165
291 Arg Gly Ser His His His His His Gly Ser Gly Ser Gly Gly His
292 5 10 15
294 atg gct gaa gct ggt atc acc ggc acc tgg tac aac cag ctg gga tcc 213
295 Met Ala Glu Ala Gly Ile Thr Gly Thr Trp Tyr Asn Gln Leu Gly Ser
296 20 25 30
298 acc ttc atc gtt acc gct ggt gct gac ggt gct ctg acc ggt acc tac 261
299 Thr Phe Ile Val Thr Ala Gly Ala Asp Gly Ala Leu Thr Gly Thr Tyr
300 35 40 45
302 gaa tcc gct gtt ggt aac gct gaa tct aga tac gtt ctg acc ggt cgt 309
303 Glu Ser Ala Val Gly Asn Ala Glu Ser Arg Tyr Val Leu Thr Gly Arg
304 50 55 60 65
306 tac gac tcc gct ccg gct acc gac ggt tcc gga acc gct ctg ggt tgg 357
307 Tyr Asp Ser Ala Pro Ala Thr Asp Gly Ser Gly Thr Ala Leu Gly Trp
308 70 75 80
310 acc gtt gct tgg aaa aac aac tac cgt aac gct cac tcc gct acc acc 405
311 Thr Val Ala Trp Lys Asn Asn Tyr Arg Asn Ala His Ser Ala Thr Thr

```

## RAW SEQUENCE LISTING

DATE: 01/27/2005

PATENT APPLICATION: US/10/671,995A

TIME: 14:52:47

Input Set : A:\BHPat.txt

Output Set: N:\CRF4\01272005\J671995A.raw

312		85		90		95		
314	tgg tct ggc	cag tac gtt ggt ggt gct gaa gct cgt atc aac acc cag	453					
315	Trp Ser Gly	Gln Tyr Val Gly Gly Ala Glu Ala Arg Ile Asn Thr Gln						
316		100	105	110				
318	tgg ttg ttg	acc tcc ggc acc acc gaa gct aac gcg tgg aaa tcc acc	501					
319	Trp Leu Leu	Thr Ser Gly Thr Thr Glu Ala Asn Ala Trp Lys Ser Thr						
320		115	120	125				
322	ctg gtt ggt	cac gac acc ttc acc aaa gtt tcg agc tct ggt agt ggc	549					
323	Leu Val Gly	His Asp Thr Phe Thr Lys Val Ser Ser Ser Gly Ser Gly						
324	130	135	140	145				
326	aat ggt cat	atg cat gga aaa act cag gca acc agc ggg act atc cag	597					
327	Asn Gly His	Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln						
328		150	155	160				
330	agc atg cat	gga aaa act cag gca acc agc ggg act atc cag agc atg	645					
331	Ser Met His	Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met						
332		165	170	175				
334	cat gga aaa	act cag gca acc agc ggg act atc cag agc atg cat gga	693					
335	His Gly Lys	Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly						
336		180	185	190				
338	aaa act cag	gca acc agc ggg act atc cag agc atg cat gga aaa act	741					
339	Lys Thr Gln	Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr						
340		195	200	205				
342	cag gca acc	agc ggg act atc cag agc atg cat gga aaa act cag gca	789					
343	Gln Ala Thr	Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala						
344	210	215	220	225				
346	acc agc ggg	act atc cag agc atg cat gga aaa att cag gca acc agc	837					
347	Thr Ser Gly	Thr Ile Gln Ser Met His Gly Lys Ile Gln Ala Thr Ser						
348		230	235	240				
350	ggg act atc	cag agc atg cat gct ctg tcc ctc gag ggt ccg	879					
351	Gly Thr Ile	Gln Ser Met His Ala Leu Ser Leu Glu Gly Pro						
352		245	250	255				
354	taataagctt	aattagctga gcttggactc ctgttgatag atccagtaat gacctcagaa	939					
356	ctccatctgg	atttgttcag aacgctcggg tgccgcggg cgttttttat tgggtgagaat	999					
358	ccaagctagc	ttggcgagat ttccaggagc taaggaagct aaaatggaga aaaaaatcac	1059					
360	tgatataacc	accgttgata tatcccaatg gcatcgtaaa gaacattttg aggcatttca	1119					
362	gtcagttgct	caatgtacct ataaccagac cgttcagctg gatattacgg cctttttaaa	1179					
364	gaccgtaaag	aaaaataagc acaagtttta tccggccttt attcacattc ttgcccgcct	1239					
366	gatgaatgct	catccggaat ttcgtatggc aatgaaagac ggtgagctgg tgatatggga	1299					
368	tagtgttcac	ccttgttaca ccgttttcca tgagcaaact gaaacgtttt catcgctctg	1359					
370	gagtgaatac	cacgacgatt tccggcagtt tctacacata tattcgcaag atgtggcgtg	1419					
372	ttacggtgaa	aacctggcct atttccctaa agggtttatt gagaatatgt ttttcgtctc	1479					
374	agccaatccc	tggttgagtt tcaccagttt tgatttaaac gtggccaata tggacaactt	1539					
376	cttcgcccc	gttttcacca tgggcaaata ttatacgcaa ggcgacaagg tgctgatgcc	1599					
378	gctggcgatt	caggttcac atgccgtttg tgatggcttc catgtcggca gaatgcttaa	1659					
380	tgaattacaa	cagtactgcg atgagtggca gggcgggcg taattttttt aaggcagtta	1719					
382	ttggtgcct	taaacgcctg gggtaatgac tctctagctt gaggcacaa ataaaacgaa	1779					
384	aggctcagtc	gaaagactgg gcctttcgtt ttatctgttg tttgtcgggtg aacgctctcc	1839					
386	tgagtaggac	aaatccgccc tctagattac gtgcagtcga tgataagctg tcaaacatga	1899					
388	gaattgtgcc	taatgagtga gctaacttac attaatgctg ttgcgctcac tgcccgttt	1959					

**VERIFICATION SUMMARY**

DATE: 01/27/2005

PATENT APPLICATION: US/10/671,995A

TIME: 14:52:48

Input Set : A:\BHPat.txt

Output Set: N:\CRF4\01272005\J671995A.raw